Approximation stability and perturbation resilience.  (19 April 2021 until May 3.)

Relevant material can be found in Chapters 5 and 6 of the BWCA book (references in questions refer to that book).

Homework assignment (hand in by May 18):

1. (Based on exercise 6.3, page 139.) For 2-median clustering, give an example of a set of points satisfying $\left(\frac{7}{5}, 0\right)$ approximation stability, but not $\left(\frac{8}{5}, \frac{3}{10}\right)$ approximation stability. For what value of $\gamma \geq 1$ is your example $\gamma$-perturbation resilient?

2. For 2-median clustering (and $\alpha > 0$, $\epsilon > 0$ of your choice), give an example of a set of points satisfying $(1 + \alpha, \epsilon)$ approximation stability, for which there is a solution that is $\epsilon$-close to the optimal solution, but its value does not approximate the value of the optimal solution within $1 + \alpha$. Are there $(1 + \alpha, \epsilon)$ approximation stable instances for which the algorithm described in class (for the case of large clusters) returns a solution that is $\epsilon$-close to the optimal solution, but not an $(1 + \alpha)$-approximation to its value? (You may assume that the value of $d_{crit} = \frac{aw_{avg}}{5e}$ is known when choosing $\tau = 2d_{crit}$ for the threshold graph, and that for each of the output clusters, the optimal center for it is computed.)

3. Recall that in class we showed a reduction from max $k$-coverage to $k$-median, implying hardness of approximating $k$-median. Prove that there is some $\epsilon > 0$ for which finding a solution that is $\epsilon$-close to an optimal $k$-median solution is NP-hard. (You may assume without proof that approximating max $k$-coverage within a ratio better than $1 - \frac{1}{e}$ is NP-hard.)

4. (Exercise 5.7, page 118.) Show that there are no instances $(X, d)$ of $k$-median with $|X| > k$ and $\gamma \geq 2$ for which for every metric $\gamma$-perturbation $d'$ of $d$ there is only one set of optimal centers, and this set is the same for instance $(X, d)$. (See Section 5.5.1 for proof techniques for metric perturbation resilience.)

5. (Based on exercise 5.11, page 119.) Show that for every $\gamma > 1$ and $\rho > 1$, there is a $\gamma$-perturbation resilient instance of $k$-median for which the single-linkage clustering gives a solution of cost at least a factor of $\rho$ larger than the cost of the optimal solution.